

October 1986

Using Archive To Efficiently Control a Network

SRIVATS SAMPATH
DSO APPLICATIONS ENGINEERING

Order Number: 231476-001

INTRODUCTION

The onset of large scale software projects has generated additional needs in all levels of the development environment. The need for a sophisticated source and version control system and efficient disk management becomes particularly important as the project team grows. This need is more pronounced at the software management level, where operating the project on schedule is of prime importance. Efficient disk management includes keeping the disk free of redundant files and keeping copies of older versions somewhere other than the disk itself. In other words "ARCHIVING" all previous versions onto another storage media that is inexpensive, reliable and transportable is key. One storage media that meets all these requirements is the NDS II tape sub-system which forms an integral part of the development environment. Moreover, the actual archive process should be easy to use, preferably automatic and should not be a drain on resources. The

ability to manage mass storage devices efficiently translates into a substantial increase in productivity for everyone. Intel realizes this need and has developed a solution that is tailored towards helping the NDS-II system manager efficiently control the development project. We introduced the TAPE SUB-SYSTEM on our Network to provide an inexpensive, reliable and transportable media, and a utility called ARCHIVE to make actual disk management both user friendly and automatic.

ARCHIVE

The ARCHIVE utility performs file backup and restoration by copying files and directories to magnetic tape or other secondary storage devices. This utility is executed at an NRM console, and with its powerful set of options, it positions itself as an invaluable tool for efficient disk management.

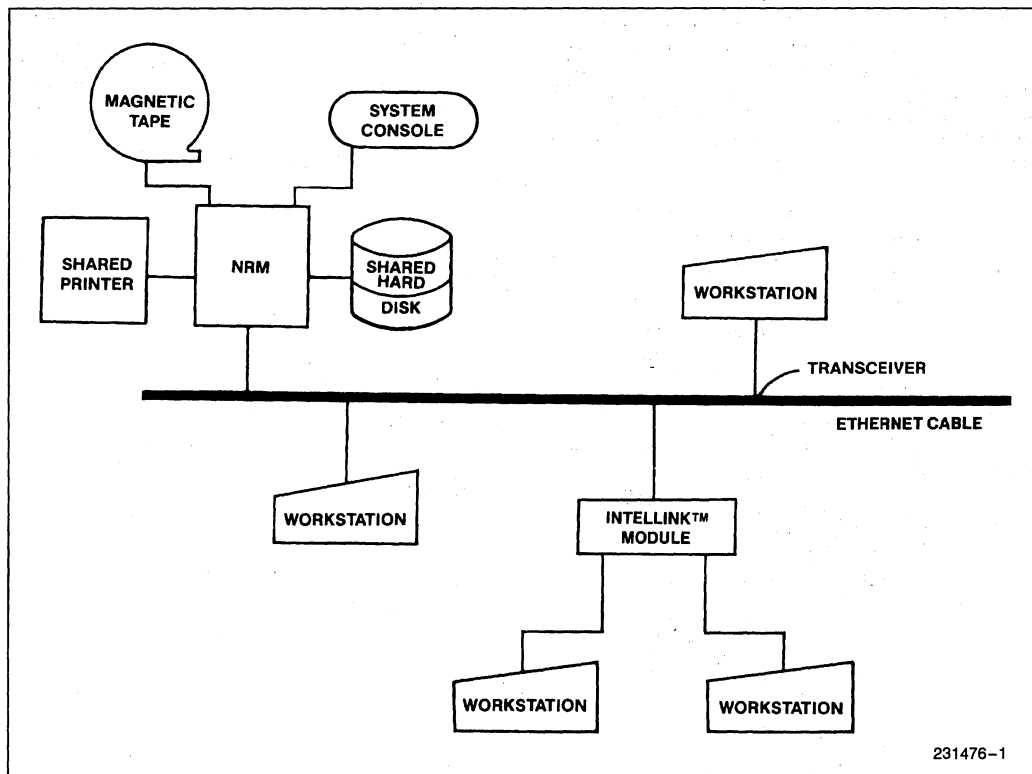


Figure 1. The Network and Its Components

WHY USE A TAPE?

Magnetic tape is regarded as the most useful storage device in the computer industry. In spite of its sequential structure, tape answers a number of requirements that can not be met by any other conventional mass storage devices. The most strong argument in favor of tape is its portability - the ability to transport tape with the minimum overhead and damage during transit makes it an extremely attractive media. Moreover, storing tapes is much more organized and efficient than storing diskettes. All these arguments lead to a single conclusion: "The Magnetic Tape should form an integral part of any development environment". The ARCHIVE utility and the tape drive together form the foundation for effective disk management, bringing about a more productive environment for any development project.

Additionally, the tape is a safety net for one of those rare disk crashes. Having backups on tape will reduce the amount of data loss in the event of a fatal disk crash. Additionally, completed projects can be saved on tape to provide more disk work space. Having these projects on tape minimizes the effort in reloading all the data if major bugs are found, or if an update is involved. Multi-project sites can benefit from the fact that tape allows easy transporting of data.

Tape backup can be classified into incremental backup and tape streamer. Tape streamer allows volume copies,

with every record on the disk copied onto tape. It is a mass data transfer from one storage device to another. This brings about a lot of overheads when only some parts of the disk need to be backed up or restored.

Incremental tape backup treats the tape as a random access device similar to a disk. Files can be added, appended or deleted, just as in a disk. This feature allows selective backup onto tape; thus eliminating the need for a mass copy operation when only a few files need to be archived or restored.

The tape drive on the NRM is an incremental backup device and ARCHIVE has been designed to use this to the fullest extent.

HOW DOES ARCHIVE HELP?

ARCHIVE has been designed to let the NDS II system manager operate the development project at maximum efficiency. Using its various options, the system manager can selectively archive files and directories onto tape, employing various qualifiers such as date accessed, created, modified, before, since, on, etc.. These qualifiers will be discussed in a later chapter (Invocation and Syntax) with examples. The options are essential for NDS-II system managers to perform selective archives of files and directories.

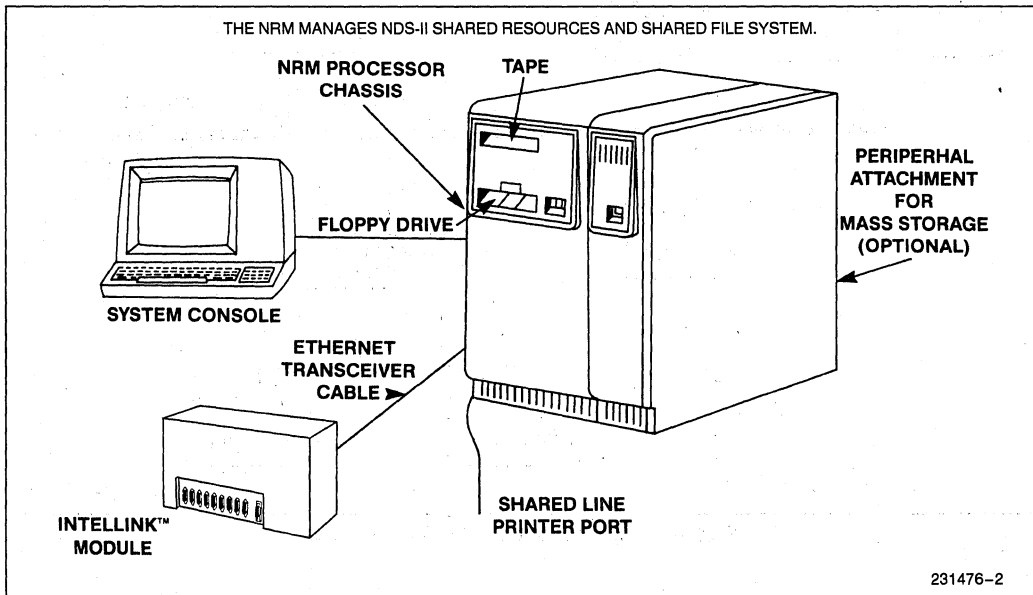


Figure 2. Network Resource Manager (NRM)

ARCHIVE can also be used in a submit or a batch file. This saves the NDS II system manager from having to type in the whole command syntax every time an ARCHIVE must be performed adding another step toward improved productivity.

Utilities like SLEEP, wakes up the system at a specified date or time, goes a long way in automating ARCHIVE. A submit file is invoked at the NRM that wakes up the system at a specified time (preferably near midnight when system load is low), archives all qualified files onto tape, then 'goes back to sleep' again. This is an important factor eliminating the need for operator intervention at any time and automating the entire process.

ARCHIVE frequency depends on the particular application and system load. It is recommended that ARCHIVE be performed at least once a week. However, in large project implementations (i.e 6 or more design engineers involved in generating or modifying more than 100K of code), ARCHIVE should be performed automatically each night. This ensures that even if a disk crash occurs, data loss is restricted to a single day's work.

All the features incorporated in ARCHIVE make it an attractive solution for effective version control and disk management. It is a productivity tool that no system manager should do without.

DATA LAYOUT ON TAPE

Tape is a sequential structured media, with all files and directories sequentially stored, but it maintains the hierarchic file structure of a disk. Every time an ARCHIVE is performed, a LOGICAL VOLUME is created, containing any number of files, from NULL to a set of files residing on a device. Each LOGICAL VOLUME has a VOLUME NUMBER and a HEADER associated with it. VOLUME NUMBERS start with 1 upwards. The HEADER is the source path name. For example:

ARCHIVE /WD0/USERS.DIR TO CTO

creates the first record onto tape and gives it the VOLUME NUMBER 1 and HEADER /WD0/USERS.DIR. Files and sub-directories under USERS.DIR will be copied in a TOP DOWN order. The same rules apply to all the subsequent sub-directories. The data layout on tape is shown in Figure 3.

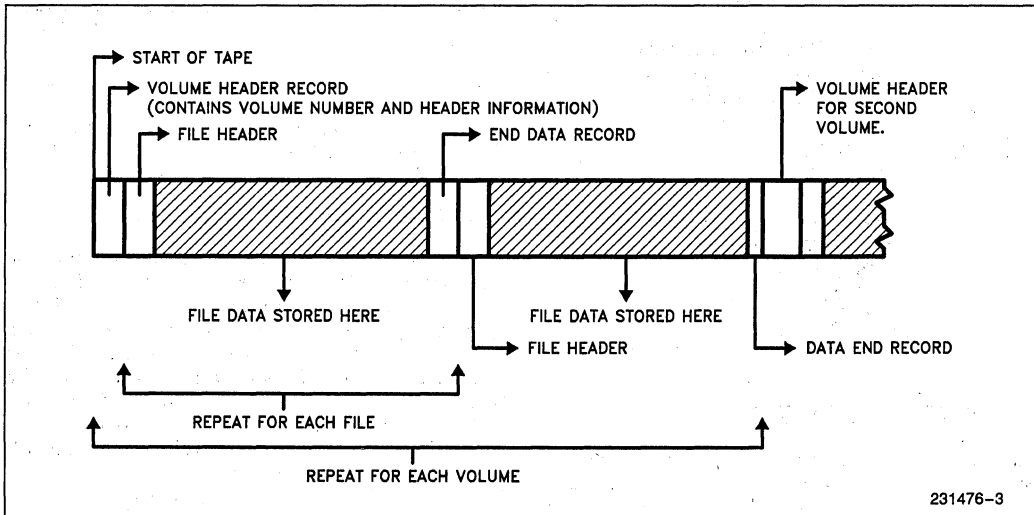


Figure 3. Format of Data on the Tape

INVOCATION AND SYNTAX

ARCHIVE, with its powerful set of options, gives the user flexibility in effectively managing the disk. The syntax of ARCHIVE is given below. During operation at the NRM, the system syntax builder prompts the user for options so none of the options have to be memorized.

The ARCHIVE syntax consists of a set of qualifiers and a set of switches. QUALIFIERS are options that qualify a file or directory for copying, allowing the user to selectively choose files and directories for archiving. SWITCHES are sets of controls that enable the user to actually control the I/O operation.

SYNTAX:

ARCHIVE source TO destination
< OPTIONS >

OPTIONS ARE:

1. QUALIFIERS:

INCLUDE, EXCLUDE (files that were ...)
ACCESSED/CREATED/MODIFIED
BEFORE / SINCE / ON
TODAY / date
hour

DIRECTORY (directory name, ...)
OWNEDBY (Owner name, ...)
FILE (path-name, ...)
AND/OR ...

2. SWITCHES: APPEND

DELETE
ERASE
LOG log-file-name
NAME physical volume name
NOUPDATE
QUERY
UPDATE
VOLUME (logical volume number)

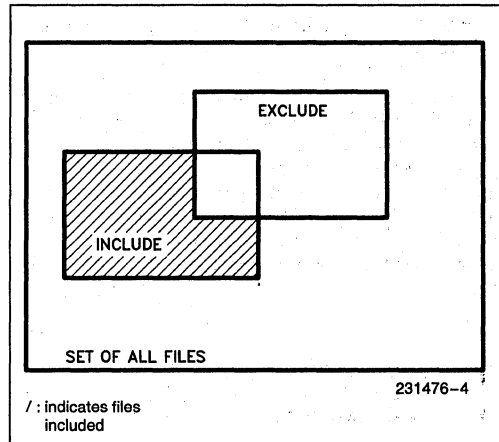
ARCHIVE QUALIFIERS

Qualifiers enable the user restrict the number of files to be archived. and discriminate against any file by time stamps (time created, modified, etc), the owner, file names and even by parent directory. These qualifiers have no limit to their length or order of appearance, an

may be specified using the keywords INCLUDE/EXCLUDE.

Include/Exclude

INCLUDE specifies the files that are to be included in the command, while EXCLUDE lists the file that can not be archived if the qualifying condition is met. EXCLUDE has precedence over INCLUDE; therefore, when both keys are used (INCLUDE files, EXCLUDE files) the following set of files will be archived:



The following is the list of all acceptable keywords for INCLUDE/EXCLUDE:

1. ACCESSED/CREATED/MODIFIED

These switches compare the time specified in the time qualifier to the last time the file was accessed, or the time it was created, or the time it was last modified. If this time agrees with the time qualifier condition, then the file is qualified. The time qualifier is required for ACCESSED and CREATED and is optional for MODIFIED. If the time is not specified with the MODIFIED switch, a default value of SINCE LAST-ARCHIVE-DATE will be used. This default value will qualify all the files which:

- Were modified since last ARCHIVE.
- Were created since last ARCHIVE.
- Were created or modified prior to last ARCHIVE but, through use of qualifiers, they were somehow EXCLUDED from being archived earlier.

Time Qualifiers:

The time qualifiers allow the user to specify an instant in time which is used in a comparison with the time a file was last accessed, created, or last modified to qualify the file for archiving:

- BEFORE Allows specifying a file accessed, created, or modified BEFORE a specific date.
- SINCE Allows specifying a file accessed, created, or modified SINCE a specific date.
- ON Allows specifying a file accessed, created, or modified ON a specific date within a 24 hour period.

— date/TODAY

For neither BEFORE, SINCE, and ON a does default date value exist. Date could be specified in two forms, either by using TODAY switch, which would read the current date from the system, or by actually specifying the date in the form of mm/dd/yy. An optional time of the day (in hours) in hh:mm:ss form with a default value of 00:00:00 can be used.

— hour

Time of the day can be specified in hh or hh:mm or hh:mm:ss. The hour qualifier is to be in parenthesis.

Examples:

```

ARCHIVE /WDO TO CTO INCLUDE CREATED BEFORE 12/21/84
;would archive all files created before DEC 21.
;a time default of 00:00:00 would be used.
ARCHIVE /WDO TO CTO EXCLUDE MODIFIED SINCE 10/10/83 ( 10:11:22 )
;archives all files, exclude those which were modified
;since 10:11:22 on October 10th, 1983.
ARCHIVE /WDO TO CTO INCLUDE ACCESSED ON TODAY , EXCLUDE & CREATED BEFORE
10/26/83 AND MODIFIED SINCE 10/24/83 ( 10:11:12 )
;archives all thefiles which were accessed today, and
;exclude those which were created before October 26th
;and were somehow modified since 11 minutes and 12 seconds
;after 10 AM on October 24th.
ARCHIVE CTO TO /WD1/DIR1 INCLUDE ACCESSED ON TODAY , EXCLUDE & CREATED
BEFORE 10/24/83 AND MODIFIED SINCE 10/26/83 ( 10:11:12 )
;all files on the tape which were last accessed today
;( exclude archive access itself ) will copied to the
;/WD1/DIR1 directory. Files which were CREATED before
;October 24,83 and were MODIFIED since October 26, 83
;will be excluded.

```

2. DIRECTORY/FILES/OWNEDBY

Allows the user to specify qualifiers other than time such as owner of files, directories, etc.

- DIRECTORY Allows the user to specify particular directories to be included or excluded in ARCHIVE. The directory could either be the full path name of the directory or partial name from where source-name left off. DIRECTORY does not accept wildcard characters. However, logical names are allowed.
- FILE Allows the user to specify particular files to be included or excluded in archive. The file name, in order to be recognized, should only be the file-name, not a path name. Wildcard characters are accepted.
- OWNEDBY Allows archive select files on the basis of owner's name.

3. AND/OR

AND/OR allows the extension of the qualifying conditions within a qualifying set. AND/OR can not be intermixed within a qualifier set defined by one INCLUDE or EXCLUDE.

4. COMMA

Comma is the separator (delimiter) between INCLUDE and EXCLUDE. In English, it makes sense to use AND in between INCLUDE and EXCLUDE. However, in ARCHIVE, you can not use anything other than Comma to separate INCLUDE/EXCLUDE. Example:

```

ARCHIVE /WDO TO CTO INCLUDE ACCESSED
ON TODAY , EXCLUDE & CREATED BEFORE
10/26/83 AND MODIFIED SINCE 10/24/83
( 10:11:12 )

```

ARCHIVE SWITCHES

SWITCHES are controls that ARCHIVE gives the user to selectively copy files and directories to/from tape or any other storage device. We will discuss each of these switches in depth to highlight the versatility of ARCHIVE.

1.0 Append, Volume

Every time ARCHIVE is issued onto tape, a Logical Volume is created. This logical volume can consist of one file or as many as all files residing on a particular device. Unless otherwise specified ARCHIVE always starts from the beginning of a tape. The tape is rewound and the header information is written followed by the

actual copy operation which copies all qualified files from the beginning of tape. Using the Append switch allows the user to have more than one volume or a related group of files on a single tape. Now instead of starting from the beginning of a tape ARCHIVE searches through the tape for the volume name specified. Default for Append is the last volume on tape.

ARCHIVE will always overwrite an existing volume if Append switch is not specified. Append to an empty tape is not valid as ARCHIVE will not know what to Append the new record to.

Recommendation: The tape should be dismounted only after ARCHIVE signs back on with the message 'ARCHIVE COMPLETE'. Use the ERASE option when writing to a new tape.

Examples: (WITH A NEW TAPE)

```
ARCHIVE /WINIO/USERS.DIR TO TAPEO APPEND
      ; This is an ERROR. No previous volume on tape
      ; to append new volume to

ARCHIVE /WINIO/USERS.DIR TO TAPEO
      ; This will create header for Volume #1 and then
      ; copies all the USERS.DIR files and sub-directories
      ; to the tape.

ARCHIVE /WINIO/SYSTEM.DIR TO TAPEO APPEND
      ; This creates a new volume on the tape (Volume # 2)
      ; and adds all SYSTEM.DIR files and sub-directories
      ; to the tape at Volume #2

ARCHIVE /WINIO/ISIS.SYS/FILES TO TAPEO APPEND VOLUME 3
      ; This skips to the third volume on tape ,writes
      ; the header for Volume #3 and then copies all
      ; files and sub-directories to Volume #3
```

2.0 Delete

This switch instructs ARCHIVE to delete all qualified files on the disk after they have been copied onto tape or disk. It is very useful when backups of older versions are performed. Once the archive process has been completed, all the old files are deleted from the source disk giving the user a better control over managing disk files. This is a disk only option.

Recommendation:

It is recommended that the user archive to tape first, using a LOG option and ascertaining that the files exist on tape. Then, he repeats the ARCHIVE to :BB: with the delete switch on to delete all the qualified files from

disk. This will eliminate any possibility of deleting files without first archiving them.

3.0 Erase

This option causes the tape to be erased before any write operation is performed. ERASE goes over all tracks on the tape and erases everything written on it. ERASE and APPEND cannot be used simultaneously, since one erases the tape and the other tries to append to non-existent volumes. ERASE is a tape option.

Recommendation:

Use the ERASE switch when archiving onto tape the first time. Use Append for subsequent logical volumes.

4.0 Log file—name

The LOG option will redirect all console messages to a specified LOG file. Errors generated because of LOG file existence will not abort ARCHIVE. (i.e. if a log file already exists it will be automatically overwritten by ARCHIVE)

Recommendation:

It is good practice to redirect console output to a LOG file when a sufficiently large ARCHIVE is being performed, keep a record of all successful archives. This LOG file should be listed and stored along with the tape.

5.0 Name physical__volume__name

The first time an ARCHIVE is issued to a tape, using the NAME option will associate the physical__volume__name with that tape. This option ensures that the right tape is used when reading from or writing to the tape. When the NAME switch is specified the name on tape will be compared against the name on the ARCHIVE command line. ARCHIVE will not continue if names do not match. The default physical__volume__name is ARCHIVE, meaning that if a NAME option was not specified during the first write operation to tape, it will be named ARCHIVE.

Recommendation:

The use of logical sounding names for the physical—volume—name of the tape is good practice. This helps in fast identification of the tape being used. Names like PROJECT1 and VERSION1.0 are good names while THIS.IS.IT and LATEST are not. The physical—volume—name should not be more than 14 characters long.

6.0 Noupdate

When archiving information from any source to a hard disk, if an existing file is encountered, NOUPDATE instructs ARCHIVE not to copy over the existing files. Thus if a file is on the disk and there is a matching file name on that tape, archive from the tape to the disk will not update the contents of the file when the NOUPDATE switch is used. This option is a default switch in submit files. If neither UPDATE nor NOUPDATE is used, the user will be queried whether the existing file should be deleted.

Recommendation:

The specified default for ARCHIVE in a submit file is NOUPDATE. However, the default for ARCHIVE in a submit file is similar to the QUERY command. If files being restored already exist, ARCHIVE will prompt the user for deletion. It is recommended that UPDATE or NOUPDATE option be specified within a submit file.

7.0 Query

This causes ARCHIVE to prompt the user for every data and directory file in the source directory, then waits for confirmation. When the user is prompted regarding a directory file, and the user chooses not to copy that directory file, none of the files and sub-directories in that directory can be archived. The default is no QUERY. QUERY used in conjunction with NOUPDATE prompts the user for every qualified file in the source directory. When confirmed that the file exists in the destination directory, the user will be informed that the file exists at the destination directory, but the file will not be copied over. QUERY used in conjunction with UPDATE prompts the user for each file in the source directory. Once confirmed, it will copy the qualified files to the destination regardless of their existence.

```

ARCHIVE    /WDO/USERS.DIR TO CTO NAME TAPE1
           ;archives every file and directory in USERS.DIR
           ;onto the tape. The tape will be named TAPE1
           ;from now on. If an attempt is made to access
           ;or write more files onto the tape with the
           ;NAME switch on, TAPE1 is the only name that will
           ;be accepted by ARCHIVE.

ARCHIVE    /WDO/USERS.DIR/MINE.DIR TO CTO NAME TAPE1 APPEND
           ;would append MINE.DIR to the tape.

ARCHIVE    /WDO/USERS.DIR/YOURS.DIR TO CTO APPEND
           ;would still work fine and appends the
           ;new directory YOURS.DIR to the tape.

ARCHIVE    /WDO/USERS.DIR/WHOSE.DIR TO CTO NAME TAPE0
           ;would be rejected with the message:
           ;RIGHT VOLUME EXPECTED.....

```


8.0 Update

This switch is the exact opposite of the NOUPDATE switch. If UPDATE is specified, all the qualified files on the tape will be copied to the destination directory, despite the previously existing files in the destination directory.

Example:

Assume that files F1 and F2 are in /W1/D1 and directory file D2 is in /W1. Also assume F2 exists at /W2/D1.

```
> ARCHIVE /W1 TO /W2 QUERY <CR>
  INDX-N11 (V2.8) ARCHIVE, V2.8
  10/26/84 11:12:33 DIRECTORY = /W1
  COPY /W1/D1 TO /W2/D1 ? Y <CR>

  DIRECTORY = /W1/D1
  COPY /W1/D1/F1 TO /W2/D1/F1 ? Y ©
  COPIED /W1/D1/F1 TO /W2/D1/F1
  COPY /W1/D1/F2 TO /W2/D1/F2 ? Y <CR>
  File Already Exists
  Pathname = /W2/D1/F2
  Delete Existing File ? Y <CR>
  COPIED /W1/D1/F2 TO /W2/D1/F2
  COPY /W1/D2 TO /W2/D2 ? N <CR>

  ARCHIVE COMPLETE

> ARCHIVE /W1 TO /W2 QUERY NOUPDATE ©
  INDX-N11 (V2.8) ARCHIVE, V2.8
  10/26/84 11:12:33

  DIRECTORY = /W1
  COPY /W1/D1 TO /W2/D1 ? Y <CR>
  DIRECTORY = /W1/D1
  COPY /W1/D1/F1 TO /W2/D1/F1 ? Y <CR>
  COPIED /W1/D1/F1 TO /W2/D1/F1
  COPY /W1/D1/F2 TO /W2/D1/F2 ? Y <CR>
  File Already Exists Pathname
  = /W2/D1/F2
  COPY /W1/D2 TO /W2/D2 ? N <CR>

  ARCHIVE COMPLETE

> ARCHIVE /W1 TO /W2 QUERY UPDATE <CR>
  INDX-N11 (V2.8) ARCHIVE, V2.8
  10/26/84 11:12:33
  DIRECTORY = /W1
  COPY /W1/D1 TO /W2/D1 ? Y <CR>
  DIRECTORY = /W1/D1
  COPY /W1/D1/F1 TO /W2/D1/F1 ? Y <CR>
  COPIED /W1/D1/F1 TO /W2/D1/F1
  COPY /W1/D1/F2 TO /W2/D1/F2 ? Y <CR>
  COPIED /W1/D1/F2 TO /W2/D1/F2
  COPY /W1/D2 TO /W2/D2 ? N <CR>

  ARCHIVE COMPLETE
```

9.0 Volume

The first time an ARCHIVE command is issued, one logical volume will be created on the tape. Subsequent ARCHIVE's to the tape using the APPEND switch create additional logical volumes on the tape. For instance, one ARCHIVE without APPEND and three more ARCHIVE's with APPEND create four logical

volumes on the tape. If the user is restoring information from the tape, not specifying volume number restores all the logical volumes on the tape. Specifying a non-existent number causes an error message and aborts the command. A valid volume number searches that particular logical volume for the qualified files and restores only files from that specific logical volume.

Example:

```

ARCHIVE  /WDO/USERS.DIR TO CTO
          ;erases the tape and copies USERS.DIR to
          ;the tape as logical volume 1.

then
ARCHIVE  /WDO/MISC.DIR TO CTO APPEND VOLUME 3
          ;causes an error message, because logical
          ;volume number 2 is not created yet.

then
ARCHIVE  /WDO/MISC.DIR TO CTO APPEND or
ARCHIVE  /WDO/MISC.DIR TO CTO APPEND VOLUME 2
          ;creates the second logical volume and
          ;copies all the files from MISC.DIR to it.

then
ARCHIVE  CTO TO /WDO/DIR1 VOLUME 3
          ;generates an error message because
          ;logical volume 3 does not exist.

then
ARCHIVE  CTO TO /WDO/DIR1 VOLUME 2
          ;copies to DIR1 all the data and directory
          ;files which were under /WDO/MISC.DIR and
          ;were archived to the tape. In a sense, the
          ;subtree starting from /WDO/MISC.DIR will
          ;be added to DIR1.
  
```

DATA RESTORATION FROM TAPE

ARCHIVE allows data restoration from tape onto disk, facilitating easy recovery from a disk crash without a significant loss of data. Reopening a project simply involves reloading all data archived onto tape. This also simplifies multi-site projects, where data can be transported and reloaded from one site to another.

In order to restore data from a tape, the user can use the device name CTO and restore all information from tape to disk. Or the user can specify a pathname to a directory on tape and restore only that directory and

associate files and sub-directories. Finally one can specify a particular VOLUME and restore information stored under that volume. Examples:

Assuming that /WD0/USERS.DIR/TEMP.DIR is empty and the tape has three records (i.e. LOGICAL VOLUMES)

APPEND and ERASE are switches that can be used only when archiving onto tape. DELETE, NOUPDATE and UPDATE switches can only be used with a disk.

```
Record #1 (Volume Number 1)
Header /WD1/USERS.DIR/TEMP1.DIR
FILES and DIRECTORIES
  /WD1/USERS.DIR/TEMP1.DIR/FILE1
  /WD1/USERS.DIR/TEMP1.DIR/FILE2
  /WD1/USERS.DIR/TEMP1.DIR/FILE.DIR/FILE3
  /WD1/USERS.DIR/TEMP1.DIR/FILE.DIR/FILE4
  /WD1/USERS.DIR/TEMP1.DIR/FILE.DIR/PASCAL.DIR/FILES5

RECORD #2 (Volume Number 2)
Header /WD0/MISC.DIR/TEMP2.DIR/TEMP3.DIR/ F1
FILES and DIRECTORIES
  /WD0/MISC.DIR/TEMP2.DIR/TEMP3.DIR/FILE1
  /WD0/MISC.DIR/TEMP2.DIR/TEMP3.DIR/FILE2

-- ARCHIVE CTO TO /WD0/USERS.DIR/TEMP.DIR

would copy all files from tape onto disk.

-- ARCHIVE CTO TO /WD0/USERS.DIR/TEMP.DIR VOLUME 2

would copy all files in VOLUME 2 to disk.
```

APPENDIX A

SLEEP:

SLEEP is a utility, available in the Network/Series IV toolbox, that executes at an NRM or SERIES IV, allowing the user to delay the execution of certain programs until a certain time. This can be included in a submit file and made to execute continuously. The sample/submit file looks like this:

Repeat

Sleep til 23:30:00

ARCHIVE /WDO/USERS.DIR TO CTO INCLUDE ACCESSED ON TODAY APPEND

End

This submit file will run forever at the NRM console and will wake up at midnight do all the archives, then go back to sleep again. Since sleep runs on the foreground at the NRM, a Cntr-C has to be performed if the user must utilize the NRM terminal for some other purpose.

This is a very useful utility in conjunction with ARCHIVE as it makes the whole process automatic and eliminates the need for operator intervention.

ACKNOWLEDGMENTS:

I would like to take this opportunity and thank Bahram Saghari in DSO Software Support for his contribution towards this Application Note. All examples on ARCHIVE were supplied by his group.